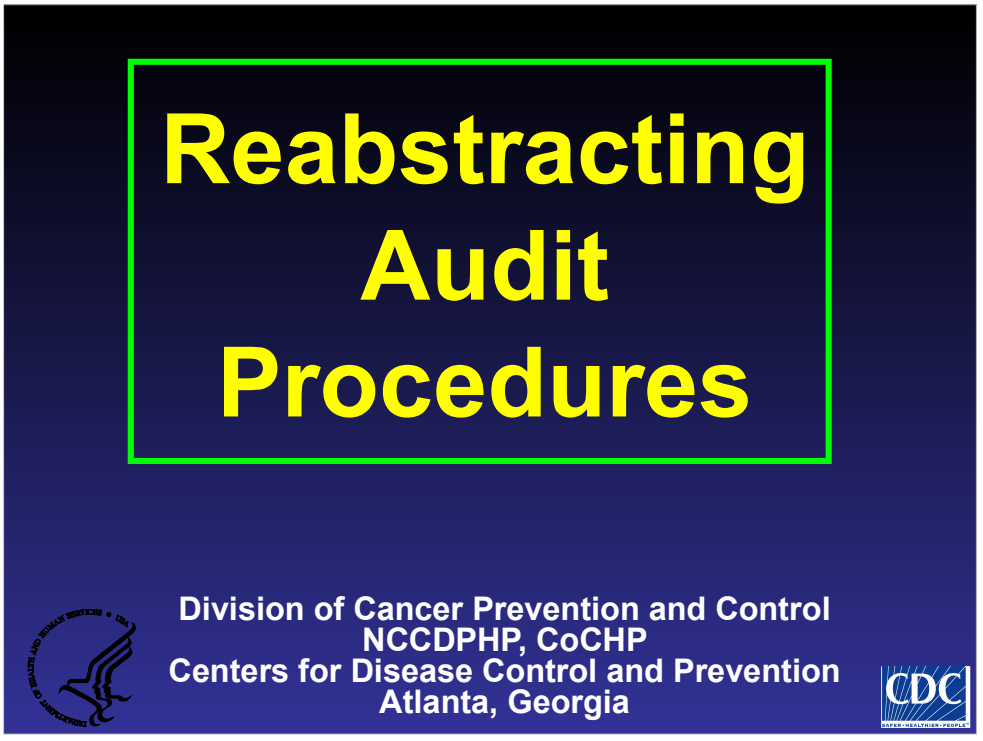
NPCR NATIONAL PROGRAM OF CANCER REGISTRIES

NPCR Education and Training Series (NETS)

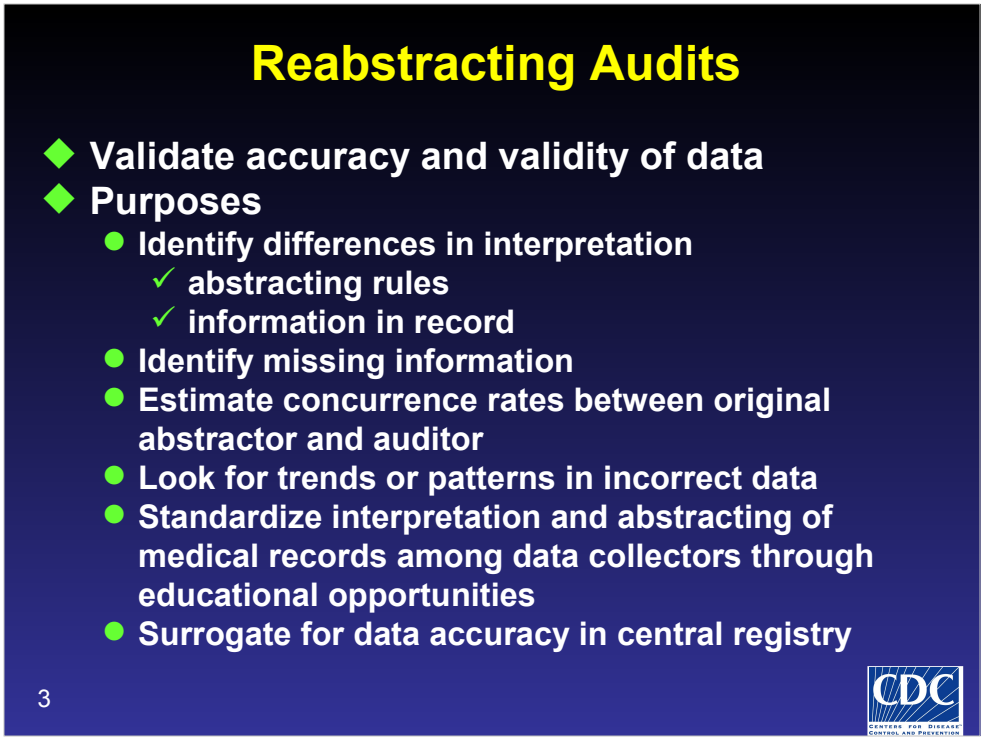
Module 3: Quality Control for Central Registries

Part 3: Reabstracting Audit Procedures

Prepared by Scientific Applications International Corporation (SAIC)
CDC Contract Number 200-2002-00576-0004
Centers for Disease Control and Prevention
National Program of Cancer Registries
Atlanta, Georgia
www.cdc.gov/cancer/npcr



NOTE TO PRESENTER: This presentation is a subset of the general quality assurance module that specifically discusses reabstracting issues and audit procedures. Some of the slides and speaker notes are derived from the general module; other slides expand on those concepts as they apply to quality control through reabstracting audits. A few comments about recoding audits are included at the end of the presentation.



A reabstracting audit compares submitted data to source documents to validate the accuracy and validity of the data. Accuracy and validity are part of the triad of quality control measures for cancer registries.

The purposes of a reabstracting audit are to—

•Identify discrepancies in the interpretation of abstracting and coding rules or in the interpretation of information available in patient records. For example, consistent disagreement between the data collector and the auditor about the coding of surgery fields may indicate that additional instructions or training are necessary in how those fields should be coded.

•Identify missing information to determine if it was missed or truly unavailable.

•Estimate concurrence or agreement rates between the original data collector and the auditor—do the abstractor and the outside auditor both arrive at the same code?

•Look for trends or patterns in incorrect data that would provide opportunities for further education and training, and ultimately, standardize the interpretation and abstracting of the medical record among data collectors through educational opportunities based on the results of the audit.

A reabastracting audit can serve as a surrogate measure of data accuracy for the central registry if the audit is performed with a sample of cases from a specific time period. Problems in the overall quality of central registry data can be identified and addressed immediately when abstracting issues are identified from a sample of cases.

Reabstracting audits are performed for a variety of reasons.

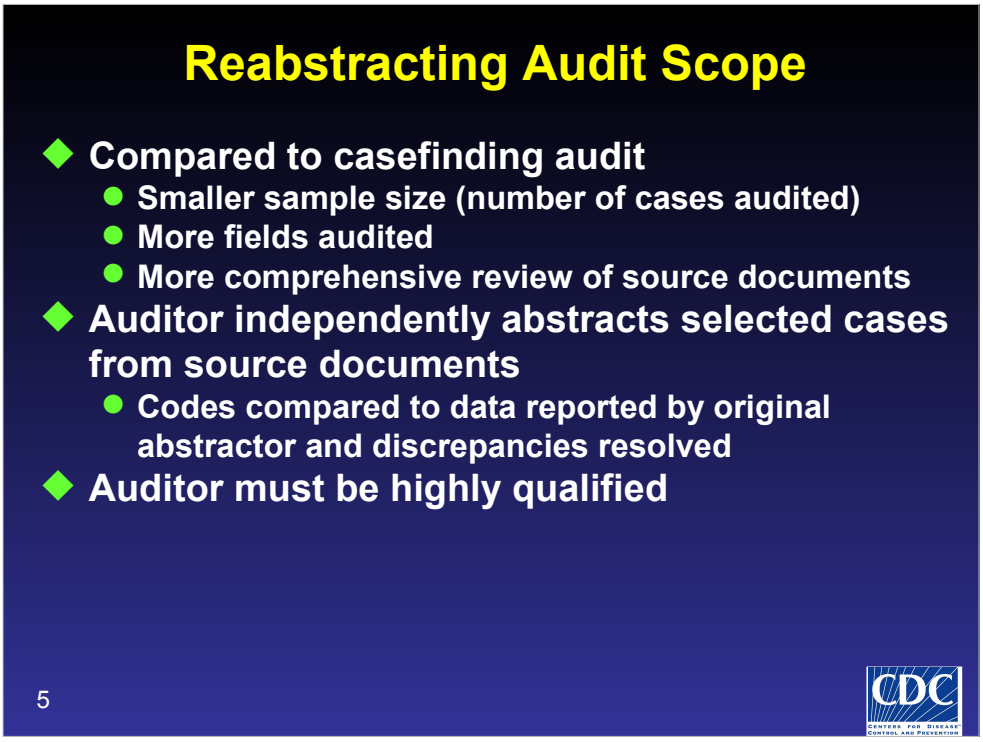
Reabstracting audits can be designed to evaluate—

- Data quality, reliability and consistency—quality is fitness for use; reliability is assurance that two people will arrive at the same code given the same information in the medical record; consistency is minimizing changes over time in how data are interpreted.
- Registry performance—are the facilities providing data to the central registry doing a good job of accurately representing the medical record in coded form.

Reabstracting audits can also identify—

- Opportunities for quality improvement: can the central registry database be made better through identification and correction of data issues?
- Training issues: are cases being miscoded for particular reasons, such as lack of education?
- Strengths and deficiencies in reporting facilities in data quality and reliability: are there systematic problems with specific types of cases being inaccurately or inconsistently coded?

Build good working relationships with reporting facilities. Most reabstracting audits are not corrective or punitive in nature, but assess understanding of rules and guidelines for abstracting and provide opportunities for abstractor education and professional development



For a reabstracting audit, the sample size is usually small (compared to a casefinding audit, for example), and the design of the audit determines which and how many data fields (such as staging or treatment fields) will be reabstracted, as that affects the number of cases that the auditor can finish in the amount of time allowed at the facility.

The procedure for a reabstracting audit is that an auditor abstracts selected cases using the case's source documents without referring to the original abstract. The auditor's codes are compared to the original abstract submitted by the reporting facility. Any discrepancies are identified in the interpretation of abstracting and coding rules and item definitions. The discrepancies go through a resolution process where the auditor and the original abstractor discuss the differences. After resolution, the remaining differences can be evaluated and reported in a variety of ways, reported back to the facility, and used in aggregated analysis of the central registry database. Training can be targeted to specific problems identified through the audit.

As with the visual editing process and other aspects of central registry data quality evaluation, reabstracting auditors must be intimately familiar with coding and abstracting rules as well as the cancer disease process and what to look for in the medical record. Previous experience as an abstractor is an important asset for a quality control staff person who performs reabstracting audits.

Developing an Audit

- Audit team
- Topic selection
- Timing
- Writing the protocol
- Sampling
- Conducting the audit
- Analyzing the data
- Feedback

6

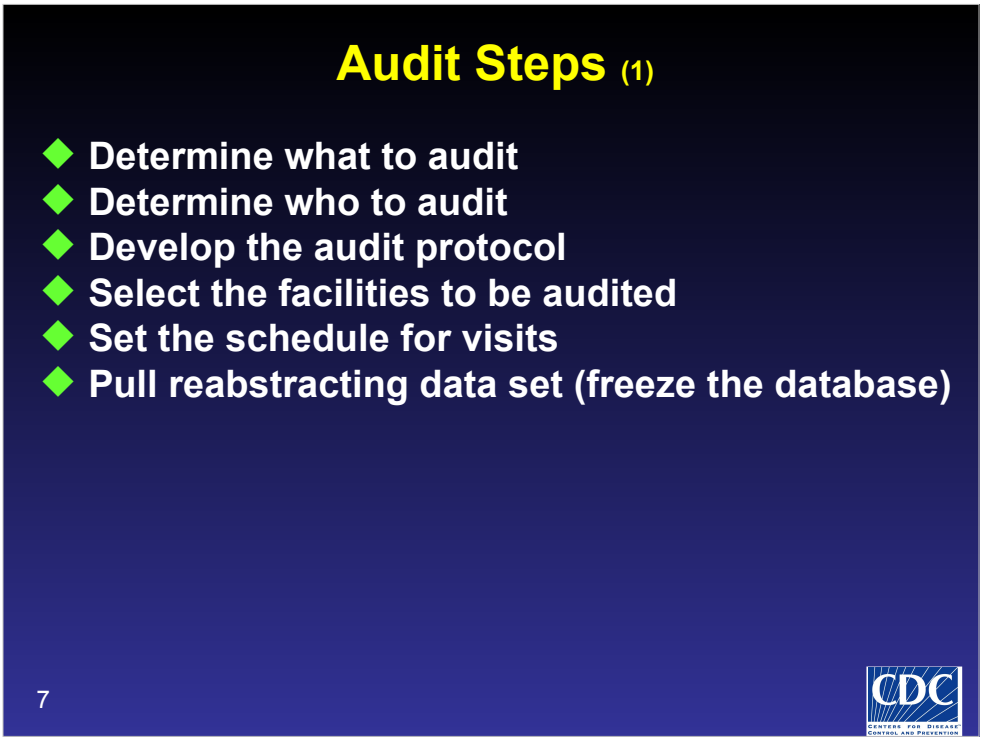
Numerous factors must be considered when any type of audit is developed. This is not a one-person job; it takes a team of central registry staff to do a good job designing and conducting the audit. There must be documentation of the process from beginning to end, including the reference documents that will be the foundation for coding rules and guidelines or case ascertainment.

There are a variety of ways to select the topic of the audit, and the timing must be carefully planned to avoid overburdening both central registry and facility staff. Sampling is both a science and a skill, and it's more than just picking how many cases to review.

The audit process begins with protocol development and doesn't conclude until the findings have been provided to the audited facilities. All of the aspects of conducting an on site audit have to be carefully planned.

The job's not finished until the data have been analyzed, summarized, and plans made for follow-through on the findings.

Let's take a look at each of these in more detail.



This is part 1 of an audit checklist—preparation for the field audit.

The first step is to determine what to audit. Many aspects of the audit must be considered, as we shall see in the next few slides. All of the audit design and development must be completed before the audit actually begins.

The next step is to determine who to audit. In many respects, this step derives from the choice of audit subject matter. In other words, once the decision is made about what to audit, the subjects of the audit will become apparent. For example, if the decision is made to audit treatment data fields, radiation oncology records should be included in the scope of the audited records.

The best way to assure consistent application of the audit design is to develop a protocol for everyone—central and facility-based—to follow.

The central registry must determine which facilities to audit. This can be done by random selection or by targeting specific facilities with case completeness issues.

Determining the actual audit schedule is the next step. The targeted facilities must be notified and requested to provide source documents.

The data set, the focus of the audit itself, must be finalized and fixed. This involves establishing a cutoff date for submitting cases in order to provide the “master file” from which the cases to be reabstracted can be randomly selected.

Basic Audit Principles

- Target the audit to get the greatest effect of available resources
- Financial considerations
  - Salaries: auditor, statisticians, programmers, support staff, facility registrars and staffs
  - Audit materials (laptops, software)
  - Controlling travel costs

8

A basic principle of developing and conducting an audit is to target the audit to get the greatest effect of available resources. This means that the central registry must use all of its resources to the maximum potential. The costs of the audit must be part of the registry's quality control budget.

Financial considerations include salaries of the audit staff, development and use of audit materials, and travel costs. The audit team consists of the auditor, statisticians, programmers, and support staff, as well as the registrar and staff of each facility being audited. The auditor must be involved in the audit planning and development processes so that he/she will know what to look for during the on-site audit, as well as the post-audit reconciliation and reporting processes.

If special software is needed for the audit, the programming costs for developing the software must be part of the budget, and development of the software must begin months in advance of the audit itself. Many states have audit software that can be modified to the needs of a new audit. If new computers are needed, these too must be budgeted, purchased, and set up well in advance of the audit. A reabstracting audit requires a reasonably current laptop with a fast processor for data entry and a fair amount of memory if the abstractor uses electronic versions of reference manuals.

The software required for a reabastracting audit is different from casefinding audit software and normal abstracting software. Abstracting software can possibly be modified to allow dual entry of data (once for the original data and once for the auditor), but the original data must be hidden while the auditor reabstracts the case. In most audits, only a limited data set is reabstracted, so the audit software must have the capability to include or exclude data fields, especially if the software is to be used year after year as audit requirements change. Additional, analyzable data fields should be included in the audit software to record discrepancy resolution comments and reasons for the discrepancy. The software should also be able to produce reports and print the reabstracted and original data fields side-by-side for use in the discrepancy resolution process. Ideally, the discrepant fields should be flagged or highlighted by the software to make the resolution process easier. Data security or encryption, back-up, and export capabilities to statistical analysis software should also be built into the software as it is being developed or modified, and the software should have the capacity to be purged once the audit is completed. Both SEER and NPCR have reabstracting software programs, as to a number of state registries.

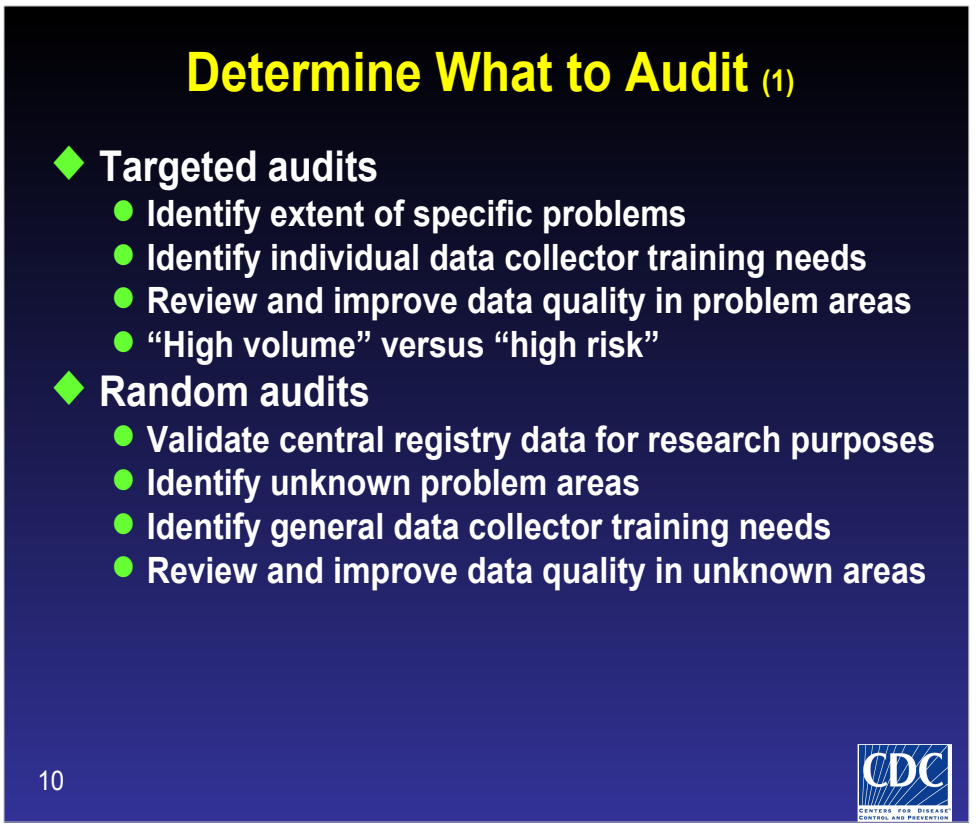
Travel costs can be controlled through creative audit procedures and extensive advance planning.

One way to control travel costs is to take advantage of electronic databases. For example, if the auditor is granted access to the facility's electronic health records, it may be possible to reabstract the cases remotely, avoiding or at least minimizing the need for travel to the site. The only limitation on remote access might be that parts of the patient's record are still on paper and not part of the electronic health record.

Another less costly—but potentially less accurate—method of controlling costs is to conduct a mail-in study, in which copies of the medical records for the selected cases are sent to the central registry. Although this sounds like a good concept, it is costly to the facility in terms of copying expenses (staff time and copying costs) and is subject to errors of omission if a copy clerk fails to include vital information in the record packet sent to the central registry. The facility may also have a policy of not releasing protected health information to anyone outside the facility, even though the central registry is exempted from privacy rules when it collects cancer data from records while at the facility.

If remote access to documents is not possible, careful planning of travel can reduce the stress on the auditor. In large states, facilities can be grouped regionally so that the auditor might be able to work from a single base of operations (one hotel) for the duration of audits in that area, or the hospital visits can be sequenced so that travel time from one day's location to the next can be minimized.

In very large states—not necessarily population-wise but in sheer distances between facilities—the central registry might send different auditors to the various regions of the state. This is particularly useful if the quality assurance system of the population-based registry is decentralized to the extent that the auditors and quality assurance personnel live in different areas of the state. However, for the sake of consistency among auditors and therefore the results of the audit, it is very important that multiple auditors receive uniform training on rules, guidelines and other abstracting principles that are subject to individual interpretation.

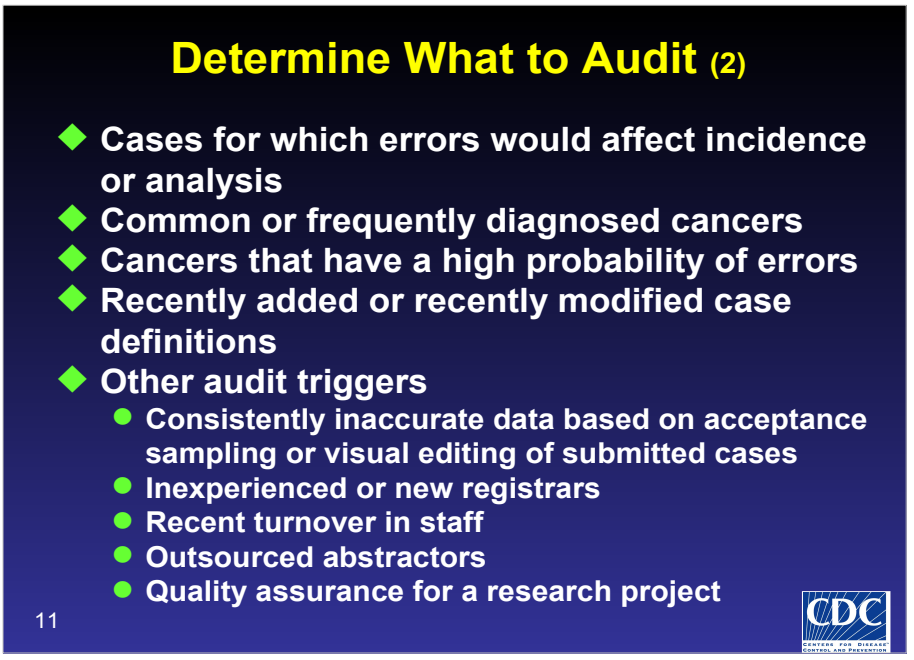


What will be the topic of the audit? This is a major decision that affects many aspects of study design. In fact, the topic may be the first aspect selected. The two big categories of audits are targeted and random. The purposes of these types of audits are different. Targeted audits have been triggered by something—usually an actual or perceived problem with the central registry data. The targeted audit will help determine the extent of the problem. Targeted audits can also identify training needs for individual data collectors. The result of the audit will be database improvement specific to the problem area, but the training or other feedback from the audit may have wider beneficial effects because of the attention paid to the target facility and data collectors.

An important concept in quality assurance methodology is to target areas of high volume or high risk for periodic audits. High-volume areas are those where there are many cases, such as the major cancer sites. Any issues identified and corrected in a high-volume area will improve a large sector of the central registry database. For example, identifying problem patterns in, and better training about, the relationships of the CS Lymph Nodes and Site-Specific Factors fields for breast will result in better quality data for thousands of breast cases. High-risk areas are those prone to error but do not necessarily involve large numbers of cases. For example, hematopoietic diseases are notoriously difficult to abstract, so a reabstracting audit that looks at coding of the morphology, staging of the disease (if applicable), and documentation of treatment would be a useful way to assess the overall quality of the central registry database. Some issues are both high-volume and high-risk, such as accurate staging of lung cancer or prostate cancer.

On the other hand, random audits aren't exactly a fishing expedition. Consider an audit of random data as a spot check or sampling of the data. It is possible that a random audit will identify problem areas that were not previously suspected and that have not been identified through formal data quality monitoring. Another use of random audits is to obtain a clearer perspective of training needs for data collectors in general, not just those with identified problems. A random audit isn't truly random; it still must be carefully planned and executed to be meaningful.

Experience has shown that because of limited resources, most audits performed by central registries are targeted in one way or another.



Basic principles of reabstracting audits include—

•Audit those cases for which errors would affect incidence or analysis. This means that the design of the audit should be based on an overall understanding of the central registry database and knowledge of what types of data discrepancies have the most potential to affect incidence counts. In other words, an audit is more likely to be productive if it targets certain types of cases or certain data items rather than performing a cursory or superficial review of all sites and data items.

•Cancers that are frequently diagnosed are a good target for a reabstracting audit because issues identified by the audit can affect large sections of the central registry database. Breast, colon, lung, and prostate are examples of these “high-volume” sites where education based on the results of a reabstracting audit can substantially improve the overall quality of the data.

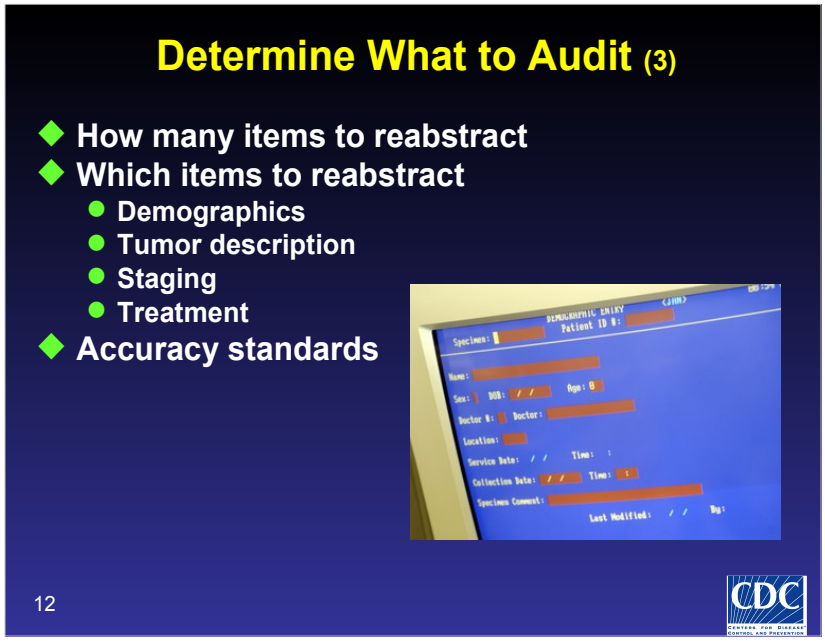
•The corollary to the second principle is to audit those cancers that have a high probability of errors, in other words, difficult cancer sites such as head and neck or lung. These “high-risk” targets may not be the most common sites, but they could be popular subjects for research; therefore it is important for both the central registry and the researchers to be assured that the data be accurate to avoid incorrect conclusions by the researchers. Specific data fields that are prone to error are collaborative staging data, dates of diagnosis and treatment, and other subjective data items.

•One concept of a reabstracting audit is to identify issues while they are still fairly easy to correct. For example, a reabstracting audit may target newly reportable cases, or those with recently modified definitions, such as the hematologic cancers that became reportable in 2001 or the and non-malignant brain and CNS tumors that became reportable in 2004.

•Any type of rule change or guideline that has recently been implemented is an ideal target for a reabstracting study. If the reabstracting audit is conducted fairly close to the implementation date, additional education and corrective measures can be identified before too many cases have been incorrectly abstracted.

Alternatively, the central registry may decide to audit a facility for other reasons. For example, the acceptance sampling procedures or process controls for a specific facility may have identified a problem in data quality coming from a single hospital. The trigger might be an increase in the proportion of unknown or blank values in key data fields. Consistently inaccurate data identified during central registry visual editing is another reason for initiating a reabstracting audit of an individual facility. Other targets might be inexperienced registrars, hospitals with recent turnover, or facilities using outsourced abstractors. Anticipation of a major research project using central registry data may trigger a preliminary assessment of data accuracy and validity.

For whatever reason, the central registry determines that a site visit and audit are necessary.



As a function of what triggered the audit, the audit team must determine what data fields to reabstract. The choices can be a very limited data set or all data items, or somewhere in between. In general, the more items that are abstracted, especially those that are not relevant to the purpose of the audit, the longer it will take the auditor to finish the case. It is better to agree on a carefully defined set of data items during the development of the audit protocol. For example, if the audit is being done to assess how well abstractors have understood and applied recent changes to coding rules, only those data fields affected by the rule changes should be reabstracted. On the other hand, if the audit is triggered by the desire to check the accuracy of a new abstractor's work, the central registry may want to review all data items. The choice of data items for the audit may be determined by how the data will be used.

Again, the purpose of the audit usually determines what data items are reabstracted, but the central registry may opt to include all items in a data category (sometimes called a data cluster), such as demographics, tumor description, staging, and/or treatment.

•**Demographics**: age, sex, race, ethnicity, and county of residence are all important factors in incidence reporting. Other demographic information such as date of birth, address at diagnosis, and Social Security number are important for case and tumor matching. Demographic information is usually very objective—the information is either there or not there in the medical record.

•**Tumor description**: primary site, histology, date of diagnosis, sequence of tumor, class of case, and tumor behavior can also affect incidence reporting, but they are equally important for researchers. Tumor information is usually fairly objective, but more subject to interpretation by the abstractor than demographic information.

•**Staging information**: the collaborative staging fields, with or without site-specific factors and the eval fields, are the way the abstractor captures the facts about the cancer and are the basis for mapping into summary stage for the epidemiologists and into AJCC/TNM stage for the clinical researchers. Staging information is subject to interpretation both by the clinician and the abstractor and must be evaluated carefully.

•**Treatment information**: date of treatment, type of surgery, type of radiation and systemic therapy are the primary data fields, but a reabstracting audit may target the reason why a particular (expected) type of treatment was not administered. Treatment information may be subjective, especially determining first course versus subsequent treatment or whether a treatment regimen is complete, but it is less open to interpretation than staging information.

Part of the audit design is predetermining accuracy standards. At present, there are no published national standards for data accuracy. Some central registries have established accuracy thresholds over time, for example, the California Cancer Registry's 97% accuracy expectations for 40 data items that are visually edited. The threshold for data accuracy can vary among data fields based on the relative importance of the field to incidence reporting or a specific research study. For example, sex and race are critical elements of incidence reporting; the threshold for these might be 100% accuracy. Tumor grade (when reported) is less critical to most research and might have an 80% or 85% expected accuracy rate. The threshold for data accuracy can even vary from audit to audit.

Where to Look

- Health information record (medical record) (all hospital encounters if filed separately)
- Pathology reports
- Radiation therapy summaries
- Outpatient records (if filed separately)

13

Reabstracting is an assessment of all the information about the case that is available anywhere in the facility. The auditor should have access to the same records that the original abstract used. This includes the health information record—or records if the inpatient records are not filed together, pathology reports, radiation oncology reports, and outpatient clinic records if filed separately from the inpatient records.

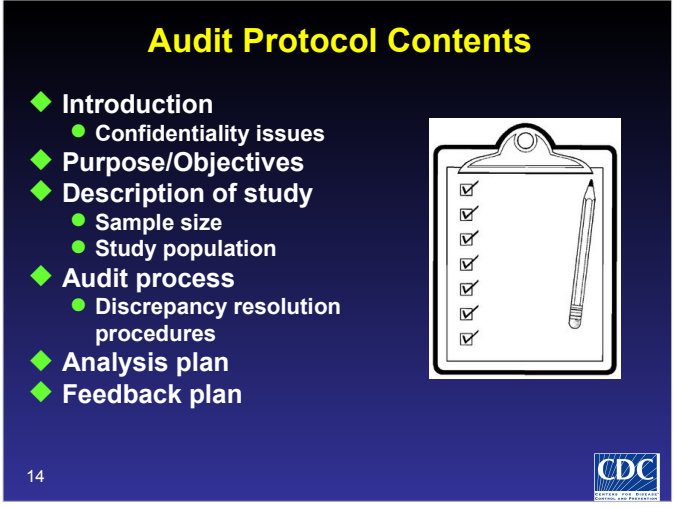
It should be the responsibility of the facility being audited to gather all possible records on each case selected for reabstracting, not just what is filed in the health information department.

The primary source of information about the case is usually the health information record, which contains demographic, staging, and treatment data. Health information departments may store their records in different ways, so it is important to request *all* of the patient's medical records, not just the most recent ones.

If the pathology report is not found in the health information record, it should be requested from the pathology department. The pathology report can provide important information about the number of primary cancers, the histologic diagnosis and grade of tumor, and the stage of the cancer at the time of diagnosis.

Radiation oncology reports provide yet another source of cancer information—and sometimes the only source of information about a randomly selected case reported by the facility. The good news is that the documentation by radiation oncologists is usually very complete (and a favorite source of information for registrars).

If outpatient encounters are not filed with the inpatient records, it may be possible to access the hospital patient master file to determine whether there are any outpatient visits. If so, those records should be requested at the time of the audit if they are not included with the patient records pulled in advance of the auditor's visit.



Any audit must be designed carefully. The audit protocol is perhaps the most important document in the audit process, especially if more than one field staff person is doing audits. It should have clearly stated objectives, a description of the sampling plan, and an outline of what is to be looked for in the analysis. All of this should be included in a written protocol.

The protocol assures consistency among participants at the central office level and keeps the audited facilities informed of the process. A thorough audit protocol will contain several sections.

1. **Introduction:** Provides authority to conduct audit and rationale for conducting audit. Any HIPAA issues or other considerations of patient confidentiality should be addressed while the audit is being planned. Because medical records will be accessed by individuals not directly involved in patient care, confidentiality policies are extremely important. Objections to accessing the records that arise at the facility can be overcome by citing the central registry's exemption from HIPAA regulations and that quality control activities are part of that exemption. Equally as important are clear guidelines for data security. We've all heard about laptops containing thousands of names and identifiers that have been stolen from cars or hotel rooms. Imagine the potential damage if a registry's detailed cancer records with Social Security numbers, birth dates and even phone numbers are stolen. Strong data security policies can reduce that risk.

2. **Purpose/Objectives:** State the reason for the audit, for example, to audit the sites with the highest potential discrepancy rates. Indicate the expected intended outcome of the audit (i.e., tabulations, lists, database cleanup—see analysis plan below).

3. **Description of study:** Identify the type of study. This will give the facility an understanding of what you are looking to find. Describe the target population (i.e., randomly selected cases diagnosed in the last half of the most recent diagnosis year, breast cancer case with positive nodes but no record of adjuvant therapy, invasive papillary bladder cancer cases). Describe the sample size (i.e., all cases, every fifth case from a specific start date, the first 25 sequential diagnoses), case eligibility, how cases are to be selected, variation or substitution allowances, and other aspects of the audit. The study population is also a consideration. It is not efficient to audit files or cases that are too old. A reabstracting audit should be done on cases from the most recent complete year, while the medical records are still available.

4. **Audit process:** Describes when and how the audit will be conducted (i.e., onsite, via Internet, mail-in documents), and the steps involved, including what documents will be requested and audited and how any discrepancies will be resolved after the audit is completed.

5. **Analysis plan:** Describes what type of calculations will be part of the final report. If possible, provide templates of the analysis tables designed by the statistician member of the audit team. This too will give the audited facilities an idea of what the audit is looking to find.

6. **Feedback plan:** Provide in the protocol a list or description of the final documents from the audit, for example a list of the "deliverables" from the central registry. These would include an indication of whether the audited facilities will receive a final report or just results for that facility, a summary report or a detailed list of findings, and any plans for education or training based on the findings.

Although the protocol describes the process in detail, it is also a good idea to provide a checklist for the facility to follow, complete with due dates and references to specific sections of the protocol.

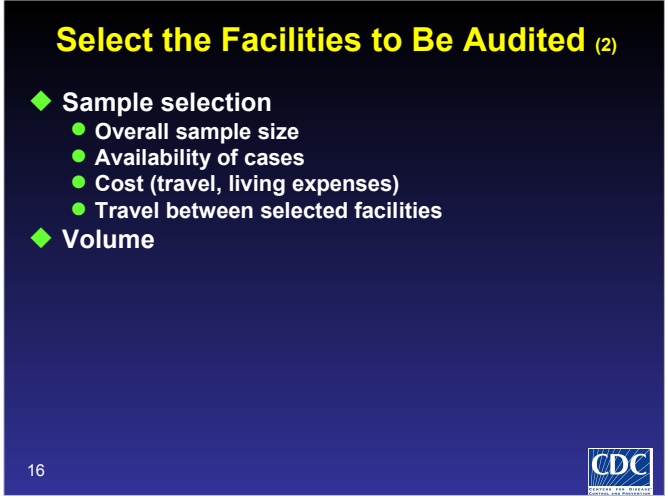
Allow as many people as possible to review the audit protocol before it is sent to the facilities being audited. All members of the central registry audit teams should read the protocol to check that, if followed, the protocol will give the desired result. In addition, it is a good idea to have one or more facilities (who may or may not be on the to-be-audited list) review the document, because there may be issues with access to records in remote storage, on microfiche, or scheduling problems that the central registry may not be aware of.

Concurrent with development of the protocol, the statisticians on the audit team should begin determining the sample selection. This is a science in itself and should not be the responsibility of the CTRs on the team. Even though the subject matter of the audit may be targeted, the selection of hospitals may still be random. If a problem is perceived in a particular site, an audit of randomly selected facilities can document the validity of the problem.

Sampling is the process of selecting cases from a population of interest (in other words, the targeted subject matter), so that it is possible to make general statements about the larger population based on the results of studying the sample cases. Samples are just that—samples. A different sample from the target population may yield a different result, so the sampling method must be carefully defined. The larger the sample, the lower the probability that the finding will have occurred simply by chance. But a large sample size must be weighed against the amount of time, effort and fiscal resources required to review the large number of cases desired.

Sampling methodologies can be very simple (random sampling) or considerably more complex (stratified, multistage, or probabilities proportional to size). In a random sample, any facility has an equal chance to be selected, regardless of facility size or location. In the facility's data base, any case that meets the criteria for the audit has an equal chance to be selected. In a stratified sample, the pool of facilities is sorted into groups either by facility size (large, medium, small) or geographic location. Then a random sample is selected from each group. In a state with a large number of facilities, multi-stage sampling may be appropriate. With this method, first the facilities are sorted into groups, such as geographic locations. A sample number of groups is determined, and then the facilities within the selected groups are randomly sampled.

A different concept called probabilities proportional to size (PPS) is used in some states and federal audits. The basic concept of PPS sampling is that the probability of selecting a hospital is proportional to its volume of cases. For example, a hospital with 1,000 cases has twice the probability of being selected as a hospital with 500 cases. The consequence, of course, is that larger hospitals tend to be audited more frequently than smaller hospital, but that's where more cases are being submitted.



Whatever the sample methodology, the number of cases to be reabstracted must be determined prior to notifying the facility. It is the job of the statisticians on the audit team to determine the total sample size for the entire audit and then how many facilities and how many cases at each facility must be reabstracted to meet the sample size. The list of cases to be reabstracted should include about 25% more cases than needed, if the facility has difficulty retrieving the cases due to readmissions, archiving, storing inactive cases, or microfilming off-site. Oversampling of cases to be reabstracted should not be an issue either for the facility or the auditor who has a quota of cases to abstract at the facility.

If a facility has not migrated to electronic health records, the availability of the cases being audited is another factor to consider. An audit of cases diagnosed many years earlier won't say much about current abstracting practices. Furthermore, some facilities microfiche or otherwise electronically store paper documents shortly after the patient is discharged. Auditing fairly recent cases can identify current data validity problems that can be corrected with educational efforts before the review and revision process becomes too extensive.

If the facility has electronic health records, it is important to plan well in advance to obtain passwords for access to the Electronic Health Record (EHR) system, or to make sure that all cases are printed out. Obviously it would be better procedure (and more cost effective) to access the records electronically in the same manner as the original abstractor.

Among the physical factors to consider when selecting the facilities to be audited, the timing of the audit and the distance between facilities are important. Take as an example a reabstracting audit for which the auditor will be "on the road" for two back-to-back 5-day periods (10 auditing days in two weeks). In larger states, consideration must be given to selecting facilities that are a reasonable driving distance apart, because the auditor will have to travel to the next location after one facility's audit has been completed for the day. The auditor's effectiveness is substantially reduced when he or she is not well rested going into the facility to reabstract. In addition, decisions must be made regarding the use of public transportation versus state-owned or privately owned vehicles. In smaller states, it may be possible to have a single base of operations for the period, with the auditor traveling from a central location each morning. Mileage and/or fuel charges, hotel expenses, meals, and smaller items such as tolls are all part of the budget planning process and also have an affect on the selection of facilities.

One more consideration is volume, or "bang for the buck." Small hospitals with fewer than 50 or 100 new cases per year may not be cost-effective to audit. Many times, there is a lower limit set for caseload, and facilities with smaller caseloads than the set limit are excluded from the sample pool. These facilities may, however, be involved in targeted audits if a problem is perceived in data quality.

Facility Considerations

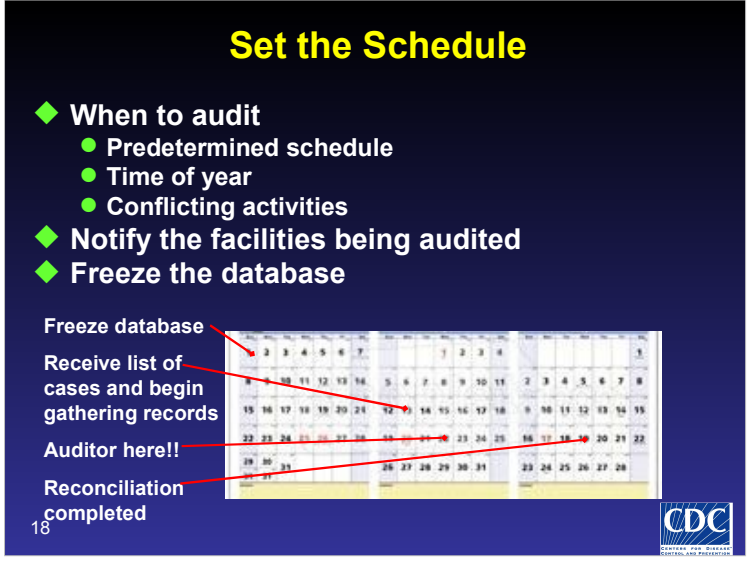
- Size and location of facility
- Available sources
- Paper vs. electronic records
- Complete audit vs. sampling

17

Sometimes the “random selection” of facilities to be audited is not entirely random.

First of all, the size and location of the facility are factors. If the reabstracting audit is performed on-site, time must be allowed for travel to the location. If the facility is small and has limited source documents (for example, no radiation therapy department), more records may be reabstracted in the course of a single day (oversampling). On the other hand, if the facility is large or the records are very complex, it might take several days to reabstract the same number of sample cases. The amount of time it takes reabstract to the medical records may depend on whether the documents are stored electronically or in hard copy, and how the hard copy records are bound. Even if the reports are stored electronically, speed of access to the documents is a factor. Also, as previously mentioned, if reports are stored electronically, it might be possible to access the files remotely and avoid the need to travel to the facility, or a copy of the file might be sent to the central registry to be worked on in the central office.

When the central registry provides the list of randomly selected cases, they should be listed in order, with instructions to retrieve the ones at the top of the list as a priority. If a record cannot be located, that case should be marked as unavailable on the sample case list and another case selected. The auditor will know from the protocol to start with the first case listed and keep working until time runs out or all cases have been reabstracted.



As previously noted, field audits are time-consuming and expensive to conduct. They disrupt the normal activities of both central registry quality control staff and the staff of the facilities being audited. Careful planning and continual communication with the facilities being audited are very important.

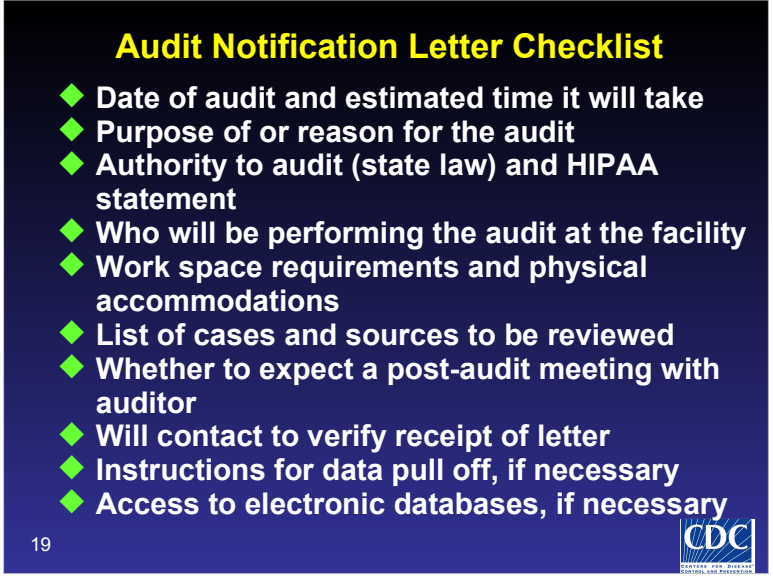
Is there a regular audit schedule defined in registry operations or standards? If so, the audit may be conducted on a pre-determined schedule—annually or more often. For example, NPCR standards require that reabstracting audits from a sampling of source documents are conducted for each hospital-based reporting facility at least once every five years. Similarly, external audits are conducted every five years by NPCR's audit contractor or periodically by the SEER Program. An individual central registry may have a policy of conducting reabstracting audits at each facility every three years. The audit schedule will most likely depend on resources such as funding for travel and data analysis and availability of staff.

An early factor in scheduling an audit must be the time of year that the onsite audit takes place. All things considered, there are only limited periods of time that an audit can take place. Usually the two to three months prior to a data submission are not good because quality control staff will be involved in final database cleanup prior to the submission. Winter months are not good choices if the auditor must drive from one facility to the next after work is completed for the day. April, May, and June are not always good choices because of the large national meetings, including NCRA, NAACCR, and the NPCR directors meeting. Having quality control staff on travel status during preparation of Request For Proposal (RFP) responses or budgets is probably not a good idea either. If the state is relatively small, it may be possible to spread out the audits, for example, doing one per week over a period of weeks or months so as not to disrupt central registry activities too much. On the other hand, if the audit planners and field staff can find one or two weeks in which to complete the audit despite a heavy travel schedule during that period, that may be a better way to get it done. Some courtesies must be given to the facilities themselves, such as avoiding Joint Commission and Commission on Cancer surveys and the period right before those surveys. Facility staff vacations and maternity leave and other factors like computer conversions or installations must be considered as well.

Once the period of the audit has been established, the facilities involved in the audit must be notified. The first communication should indicate that the facility has been selected for a reabstracting audit on such-and-such a day and it should provide a copy of the audit protocol and a checklist of preparation activities. Details of the audit notification letter will be discussed in a moment.

As part of the checklist, the facility should be given a date to submit a final database to the central registry. This is sometimes called freezing the database. Adding cases to a database is an ongoing process, so freezing the database provides a cut-off date for the central registry and a fresh, updated database from which to select the cases for reabstracting. Freezing the database means that the contents of the facility's database as of a specific day are copied into a file and submitted to the central registry. Having all facilities selected for audit freeze their database as of one specific day helps control one more variable for reabstracting.

Providing a cut-off date several weeks ahead allows the audited facility to make sure that cases in suspense are finalized and included in the frozen database.



Notifying each facility to be audited is both a courtesy and a necessity. This should be done both by e-mail (to more than one person at the facility) and by formal letter. This slide shows a list of all of the items of information that should be included in the notification letter, and the protocol should be included as part of the notification package. Almost as important as the date of visit are the work space requirements and physical accommodations needed for the auditor(s). Even minor things, such as a power outlet, should be mentioned. Too many times an auditor has had to balance the audit laptop on his or her knees because there wasn't any desk space available due to inadequate advance preparations.

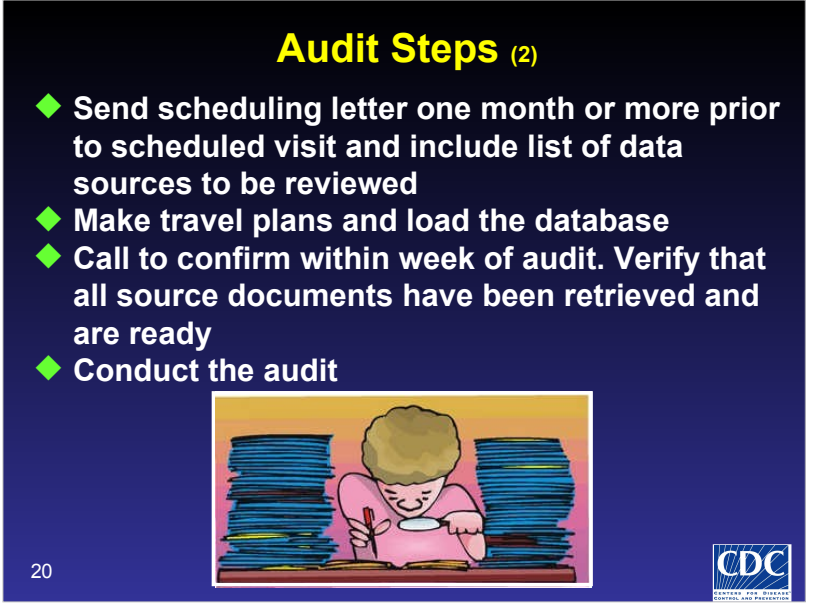
The reason for the audit, such as a routine audit or an audited triggered by some data issue, should be clearly stated in the first few paragraphs of the notification.

It may be convenient for the recipient to design the notification letter as a check-off list with due dates for data pull-off, reminders for notification of other departments (pathology, radiation therapy, outpatient file room as needed), and responses to other actions indicated in bold. It is also important for the facility to know to whom they should send the response acknowledging receipt of the notification.

For a reabstracting audit, the notification letter should include the list of randomly selected cases to be reabstracted (including the oversampled cases) and specific instructions that *all* medical records pertaining to the case should be retrieved and waiting for the auditor.

On the facility's end, any departments involved in the audit should be notified as early as possible, then again a few days in advance of the visit to retrieve any records, and once more the day before the audit. Any questions about the audit process should be directed to the central registry audit team.

One particular issue must be considered well in advance: electronic medical records. If the central registry expects the records to be printed, that will take extra time, and the clerk in charge of the print job may not get all pertinent information if the entire chart(s) are not printed. Alternatively, will the auditor expect access to the electronic record while at the facility? That may take special permission from the facility and/or a assignment of a special or limited access password.

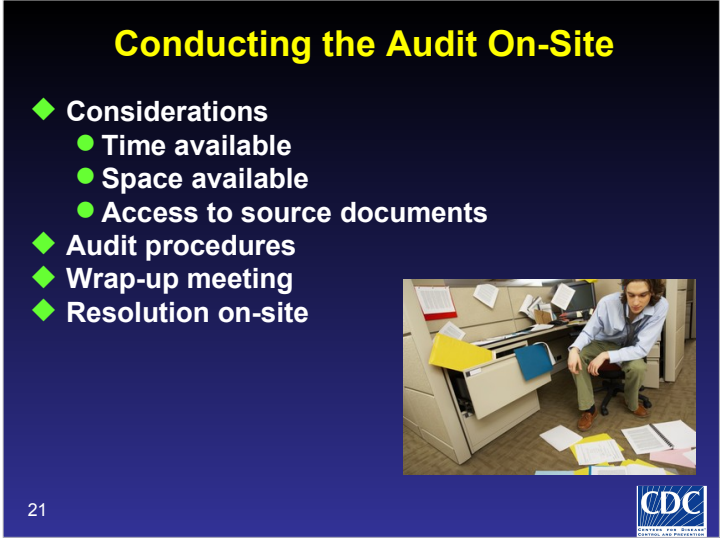


The second part of the audit process is to maintain contact with the facility regarding the specifics of the audit as the date of the audit approaches, particularly if the facility has missed any deadlines for submitting the requested databases or responding to questions. As previously noted, the facility should have a copy of the protocol and a checklist of items to have ready. If the audit protocol states that several radiation oncology charts are included in the reabstracting process, the facility must notify the radiation oncology department to pull these records. The same goes for records needed from the health information department or outpatient filing area.

Finally, the field staff can begin to make travel plans. If the entire audit is to be conducted over a 1- or 2-week period, the dates the auditor will be at a specific hospital must be confirmed. Driving directions and hotel arrangements can also be worked out. As a reminder, the auditor needs to pack or load on the computer any necessary reference materials (ICD-O-3, *FORDS* manual, state data collection manual, *Collaborative Staging Manual*, *TNM* Manual, and any other references) and learn any data backup or transmission procedures established by the central registry's IT department.

The abstracted cases selected from the facility's frozen database must be loaded into the audit software on the auditor's laptop computer. These cases contain personal health information that must be protected through both confidentiality and data security procedures.

A week before the on-site audit, contact the registrar at the facility (or other point of contact if there is no registrar) and verify that everything will be ready for the auditor—in other words, that other departments have been notified. This should include confirming that the points of contact will be present in the other departments (not on vacation), and that there will be space for the auditor to work. This is also the time to confirm local hotel accommodations, driving directions to the facility, parking information, and directions within the facility to the office of the principal point of contact, as well as any other plans the facility may have for the auditor, such as a group lunch or late afternoon wrap-up session.



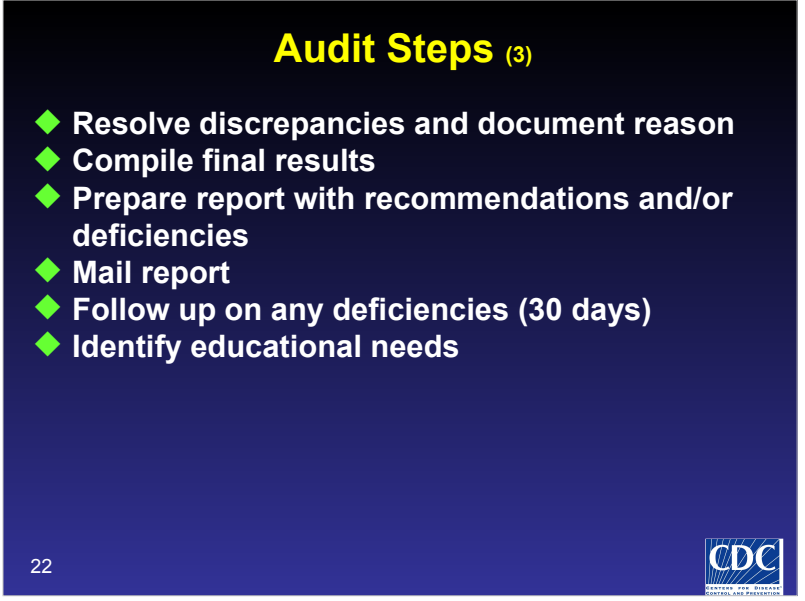
On the day of the audit, the auditor should be present at the office on time and ready to begin work. The auditor should present a business card, letter of introduction, and any documentation or confidentiality agreements that should be signed. As a courtesy, the point of contact should accompany the auditor to any of the departments where work is to be done and introduce the auditor to the staff there.

What happens during an on-site reabstracting audit? For each case selected for reabstracting, the auditor independently reviews source documents and enters codes and supporting text documentation into the reabstracting software. When the case is saved, or at the end of the day (depending on the software design), the audit program will compare the original codes it has stored for the case with the codes the auditor has entered and identify discrepancies. It may be better that the auditor not see discrepancies until the audit has been completed to avoid biasing cases abstracted later in the day.

It goes without saying that the auditor has to be flexible on-site. If the working-space arrangements are not satisfactory, a request can be made for a different location, but sometimes the space is just not available to be comfortable and efficient at the same time. If the number of records to be reviewed exceeds the time allotted, the auditor may have to stay into the evening to finish things up, or may have to resort to alternative plans (scheduling another audit day or using a smaller sample of cases). At the end of the day, all work should be backed up before leaving the facility.

The auditor may have arranged to have some sort of wrap-up meeting at the end of the day to thank the facility for their cooperation and provide some preliminary findings, if possible.

If time and software allow, the discrepancy resolution process can begin on-site. For a reabstracting audit, the auditor and the original abstractor can discuss any discrepancies in coding while the source documents are readily available. Although any decisions made at this time will not necessarily be final, beginning the resolution process while everything is still fresh in mind is one way to make the process go more efficiently.



The final part of the audit is wrapping up the paperwork after the on-site audit is completed. This involves resolving any data discrepancies with the facility data collector, tabulating the results of the audit, completing the final report, determining any recommendations, and identifying any educational needs based on the audit.

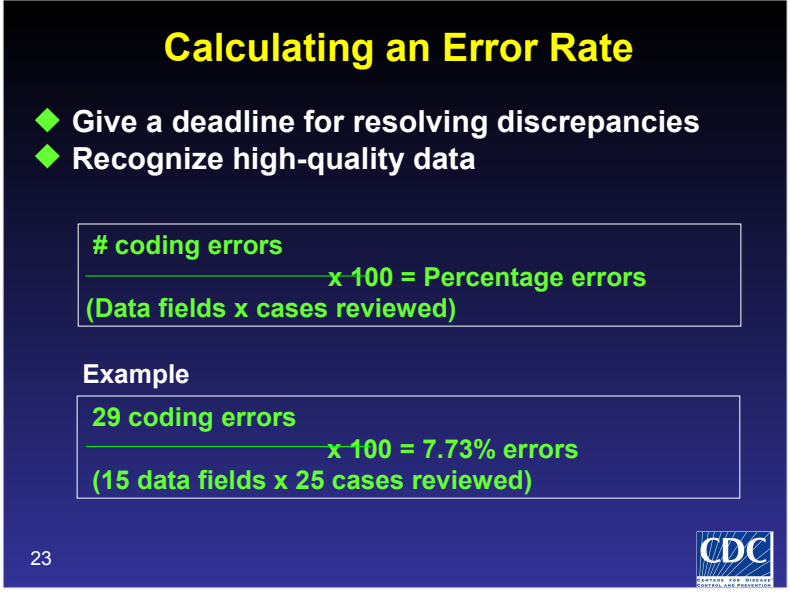
For the resolution phase of the audit, the audit software should provide a list of data fields for each case showing the auditor's codes and the original codes. On-site or through one or more conference calls, the auditor and the abstractor can work through the discrepancies and determine which is the better code for the case. During the resolution process, it would be useful for both the abstractor and the auditor to have as a reference a written checklist of related fields, similar to that used for visual editing.

The discrepancy resolution process allows the facility abstractor to review the source documents to determine if the original codes were correct or another code should have been chosen. The resolution may be information was missed or not available when the case was originally abstracted, that there was a data entry or coding error, or that coding rules were misinterpreted. On the other hand, the resolution may be that the auditor incorrectly reabstracted the data item and the original data is not in error. For each discrepancy, the reabstracting software should keep track of the decision on the correct code and a reason (preferably from a coded list) for the discrepancy.

There are many reasons for discrepant codes. At least one reason should be indicated at the time of resolution. These are some examples:

- Abstracting error, information was missed or incorrectly abstracted.
- Better information obtained from another source, such as a different facility.
- Coding error.
- Difference in registry's interpretation of coding rules.
- Data entry or computer error, key error, programmer error.
- Registry information was incomplete.
- Medical record not available.
- Related error: results from a primary site code change may cause a related error in laterality, extent of disease, or surgery fields.
- Auditor error.
- Unknown reason for difference.
- Registry information was incomplete at time of data submission.
- Both codes are wrong (auditor and original code are in error).
- Could not resolve; refer to central registry for resolution.

There will be occasions when the auditor and the original abstractor may have completely different interpretations of the information in the medical record and cannot come to a resolution. Such cases should be referred to an impartial but knowledgeable "arbitrator" at the central registry or an expert at a national registry for a decision.



The discrepancy resolution process may take multiple conference calls over several weeks if there are a lot of discrepancies. It is best to state a specific deadline for completing resolution in the audit protocol and notification letter, as well as during the wrap-up meeting. After that deadline, any discrepancies will be counted as true errors.

For a reabstracting study, there should be an overall error rate. In addition, the central registry may decide to subcategorize differences (which after resolution can be described as errors) into major, minor, and unknown-to-known. In addition, there may be contextual errors; for example, if the primary site is incorrect, all of the staging and treatment codes will be incorrect. A decision will have to be made whether that will count as a single error or the total number of incorrect data fields.

The error rate is a calculated percentage. The numerator is the number of errors. The denominator is the number of data fields reviewed per case multiplied by the total number of cases reviewed. The numerator divided by the denominator produces a decimal fraction, which, when multiplied by 100, produces an overall error rate.

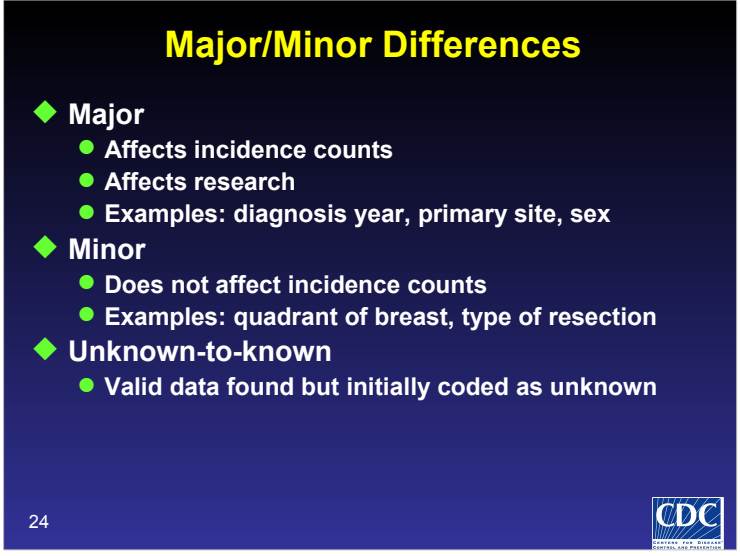
For example, at Facility A, the auditor identified 29 coding errors during the audit (the numerator). Fifteen data fields were abstracted for 25 cases, for a total of 375 possible errors (the denominator). The calculation would be 29 coding errors divided by 375 possible errors and multiplied by 100, for an overall error rate of 7.73. Whether this is good, bad, or in the middle would depend on the central registry's accuracy standards as established in the audit protocol.

The facility's abstracting accuracy rate is the corollary to the error rate, in other words, 100% minus the error rate or 92.28%. Alternatively, the accuracy rate can be calculated as the total number of correct answers divided by the total number of data items audited (data fields times cases reabstracted).

There are many more ways to analyze the results of a reabstracting audit compared to a casefinding audit. In addition to the overall error or accuracy rate, the final report might include a distribution of which data items were most commonly in error, and a summary of the types of errors based on the previously mentioned reasons for discrepancies, such as—

- Abstracting error (information missed).
- Auditor error.
- Coding error (misinterpretation of coding rules).
- Data entry error.
- New information in the medical record not available when originally abstracted.
- Other and unknown errors, including unresolved issues.

Incentives based on quality assurance audits can encourage improvements in data quality at the facility level. A wide range of incentives are used by various central registries, including letters of commendation for the registrar and copied to the facility's administrators, plaques or certificates, reference materials and access to products to improve the registry (such as death index access), and development of a data quality report card that allows reporting facilities to compare their results to others in the state.



The outcome of a reabstracting audit is an error rate or, conversely, an agreement rate. However, the definition of an error is relative to how the data will be used. Some central registries categorize errors as major or minor as part of the reabstracting audit report. For example, demographic information, particularly county/state of residence, age, sex, and race are critical measures for incidence reporting and the calculation of incidence rates. So too are primary site and tumor behavior. If these data fields can be visualized in an incidence table by primary site, any shift from one cell to another or into/out of the table itself is an important, or major error. If the incidence table involves only invasive tumors, incorrect designation of the behavior is an error. Shifting a case from lung to prostate because the case was incorrectly abstracted as a lung primary is a major difference.

Other differences may be considered major in the context of research. if a researcher is interested in adenocarcinomas of the colon, the county/state of residence may not matter, and the age, sex, and race of the patients may be of interest for descriptive statistics. The key pieces of information in such a study would be the location of the tumor in the colon, the histologic cell type, and probably the stage and treatment information. If the research involves differences in treatment by stage, any shift of stage or type of treatment for a case would be considered a major difference. If the central registry uses major and minor differences as part of its analysis, it should clearly define the context in which they are included.

In most circumstances, a minor difference does not affect incidence counts. Even in research, the subcategory site codes of most primaries are analyzed together, such as the different lobes of the lung or the areas of the bladder. It may not matter whether a surgical procedure was an excisional biopsy or complete removal of the organ.

A difference in Collaborative Staging Extension field codes could be major or minor, depending on the derived stage that results from the codes. For example, in lung, code 10 and code 30 both map to summary stage localized, so that would probably be a minor difference. But there would be a major difference between code 10 (localized) and code 45 (regional by direct extension) in a research study based on stage at diagnosis. For breast, tumor size is an important factor for mapping to the T category in TNM, so a difference in tumor size between ranges 001–020, 020–050 and > 050 would be major differences, but differences within each range would be minor.

One other audit difference is worth noting: unknown-to-known. This occurs when the original abstractor codes a data field as unknown and the auditor finds enough information to code a specific value. Unknown-to-known differences may be a sign of inattention to detail or a signal that the abstractor needs additional training in how to interpret information in the medical record.

Reabstracting Agreement Rates

Required Field	Major	Minor
Address at Dx - State	✓	
Race	✓	
Sex	✓	
Birth Date	wrong ccyy	wrong mm/dd
Date of Diagnosis (mm/dd/ccyy)	>30 days	<30 days
Primary Site	wrong site	wrong sub-site
Laterality	✓	
Histology Type	wrong histology	wrong subclass
Behavior Code	✓	
Grade		✓
Summary Stage	✓	

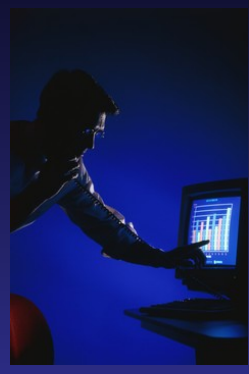
25

Most reabstracting analysis is focused on whether any identified discrepancies affect (incidence) data quality. In this context, most states that perform reabstracting studies have developed or adopted tables of major and minor differences. The examples are from a previous NPCR reabstracting audit. The most recent NPCR audit did not designate errors as major or minor. An error was simply and error.

Major differences affect incidence accuracy; minor differences affect other types of research. However, major-minor differences can be adjusted based on the planned use of the data.

Analyzing the Data

- Predetermined error rates
- Predetermined target thresholds
- Benchmarks and standards
- Calculations

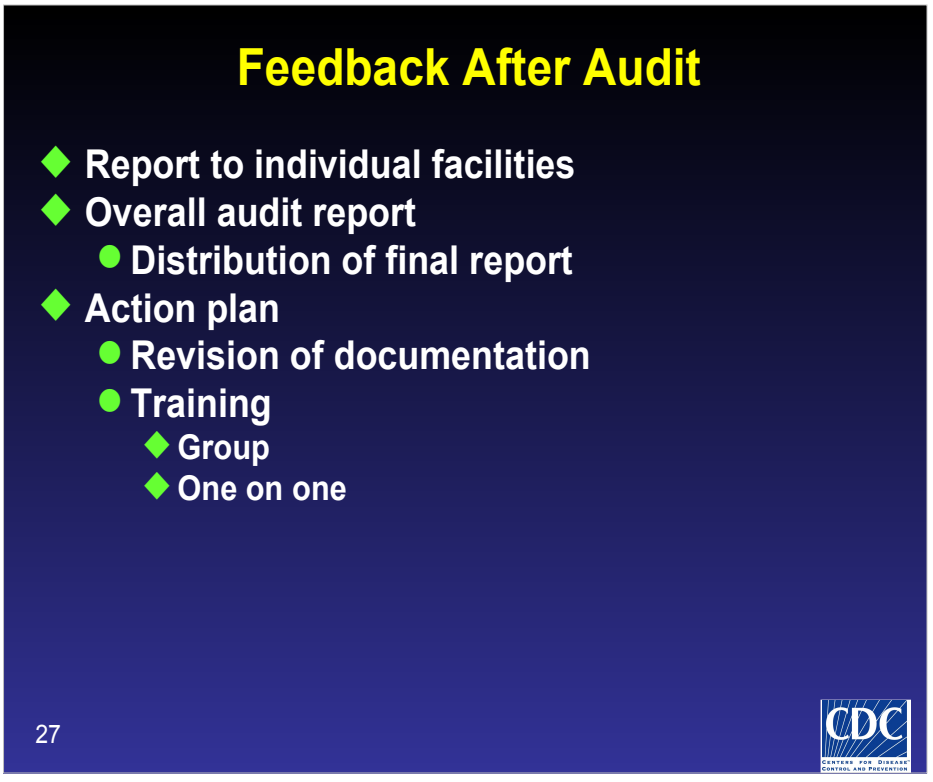


26

When the discrepancy resolution is complete, the central registry audit team can begin the analysis of findings, first by individual facility, and then in aggregate. The audit design will affect how the data are analyzed. For example, the audit team may have made a decision to compare all facilities against a standard, such as a threshold for accuracy. This might be considered a “pass-fail” type of standard—if the facility exceeds the threshold, it passes, regardless of how “perfect” it is. If the threshold is not met, the facility fails, regardless of how many errors were made. Another option might be considered a grading system with ranges for error rates on a scale of one to five stars, or A-B-C-D-F, or poor-satisfactory-excellent. With any of these options, the ranges would be predetermined as part of the study design. Either way, the central registry must decide what the acceptable quality level should be in advance of the audit itself.

Benchmarks will vary according to the type of audit and its purpose. For reabstracting, the benchmark is usually the established standard by the central registry conducting the audit, since there are no published federal standards for data accuracy. For data quality, benchmarks can also be obtained from published reports developed by other central registries and the Commission on Cancer’s National Cancer Data Base, as well as by reviewing journal articles and documents issued by NAACCR.

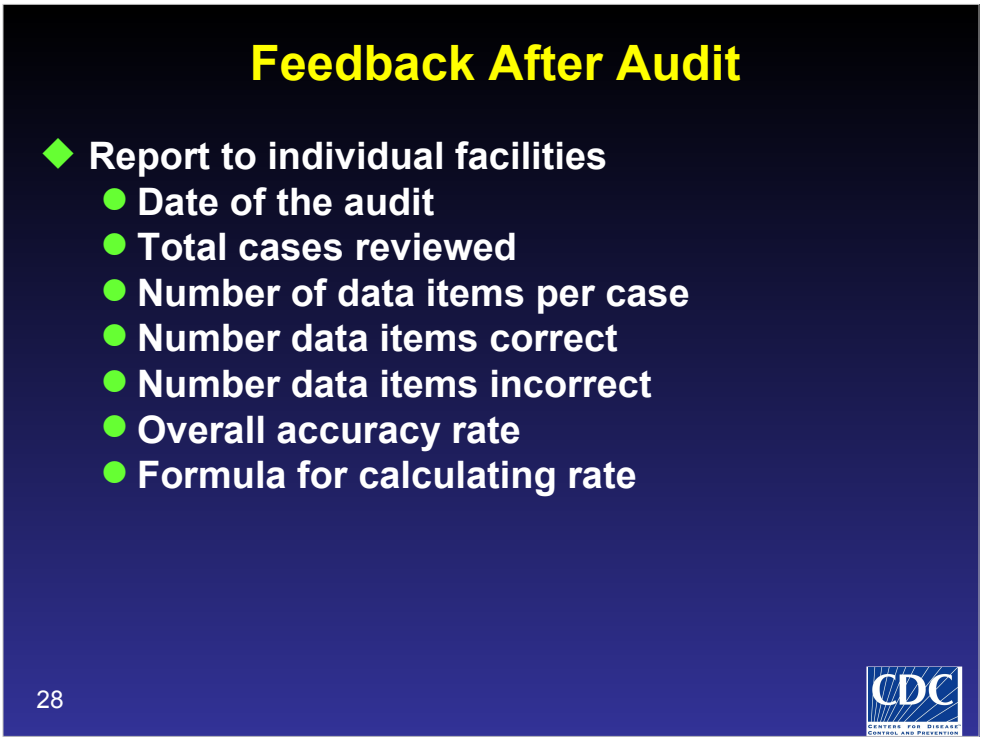
The calculations may be considerably more detailed than just a simple percent of data items determined to be “wrong.” More in-depth calculations may be reported, such as the percent of errors by data ‘cluster’ (demographics, tumor identification, staging, treatment), by type of case (inpatient, outpatient), or by primary site. If available, data from previous reabstracting studies of similar cases can be included in the current report to identify improvement or further deficiencies.



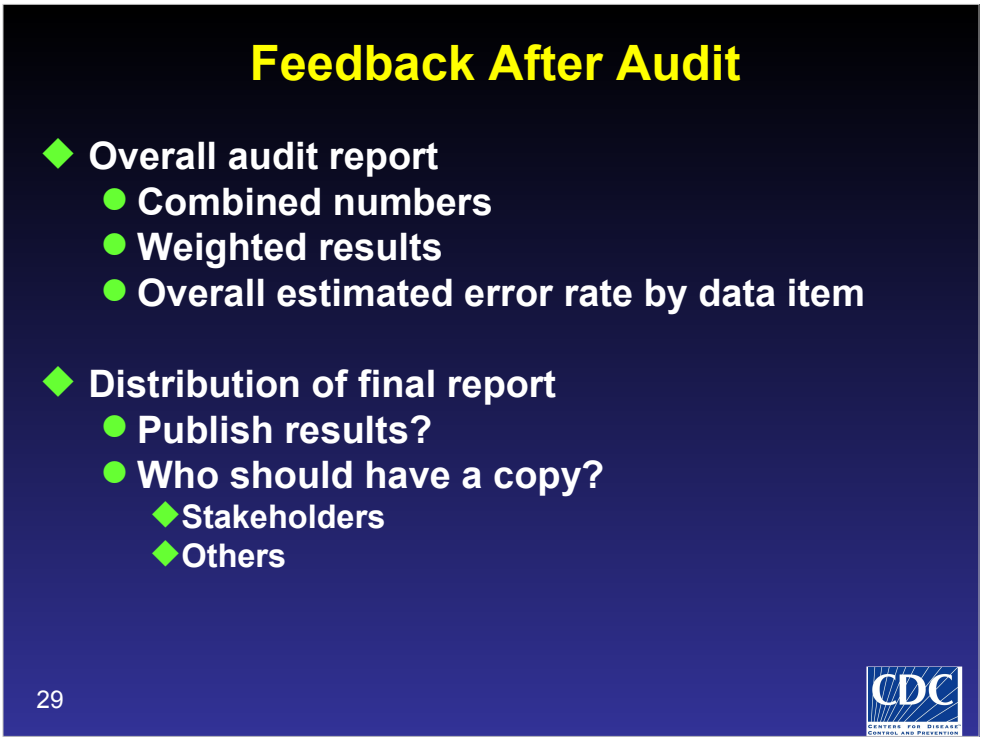
The last step in an audit is feeding the results back to the facilities audited and creating an aggregate report. This is more than just common courtesy for those who were involved in the audit, it is closing the loop on data quality assurance. Unless the facility and the central registry are notified of the findings, they will not be able to take the appropriate actions to correct the errors and develop strategies for education or process improvement that hopefully result in a lower error rate during the next audit. Eventually, there will be another audit to monitor whether the corrections were effective. Let's take a closer look at each of the bullets on this slide.

Finally, the overall report needs an action plan. The action plan is based on interpretation of the analysis and the overall experiences of the auditors. For a reabstracting audit, did the final analysis indicate that only certain fields (or groups of fields like the staging section) were problem areas? If so, the action plan should describe training efforts, either for all data collectors in the state or for individual data collectors. Did the audit reveal that there were areas in the central registry's documentation that were not clear. If so, the action plan should describe the areas of the coding manual or the reportable list and reporting requirements that should be clarified. Did the audit indicate that the protocol was not clear enough to the participants? If so, revise the protocol now while it is still fresh in mind, so that the improvements won't be forgotten the next time a similar audit is done in the future.

This reabstracting audit is complete. Now it's time to step back, take a deep breath, and begin the quality improvement cycle once again.

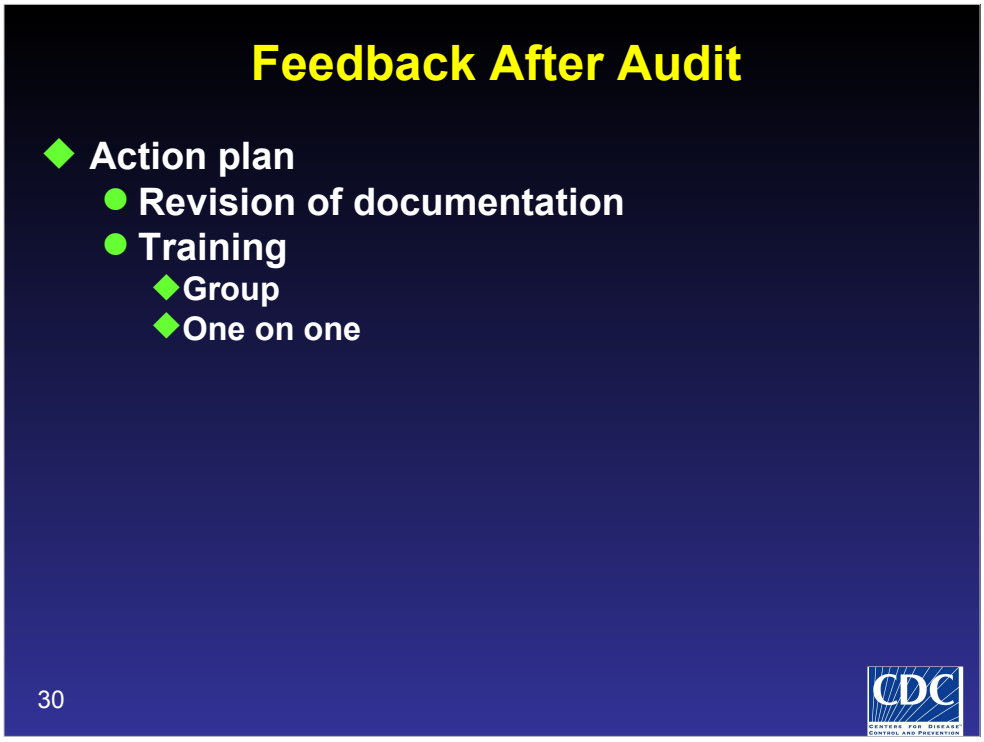


For a reabstracting audit, the report to an individual facility should include, at a minimum, the date of the audit, the total number of cases reviewed, the number of data items reviewed per case, the number of correctly abstracted data items and the number of incorrectly abstracted data items, together with an overall accuracy rate and the formula for calculating that rate. It would be helpful to summarize the errors by major and minor categories and discuss any pattern of errors. Reports to the individual facilities can be in boilerplate format, inserting facility-specific data in a generic report. A cover letter that contains thanks for cooperation and congratulations for a good job (if applicable) should be sent to the point of contact (registrar) with a copy or separate letter to the hospital administrator.



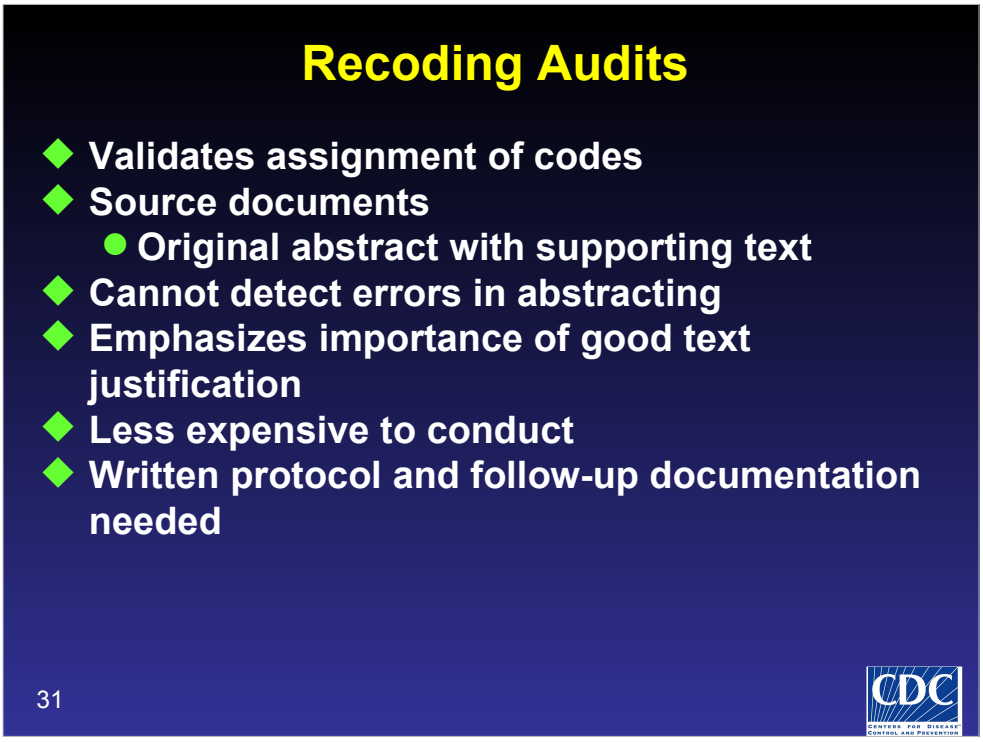
The overall audit report will be combined numbers from all the audited facilities. The statisticians on the audit team will have to determine whether the results should be weighted. For example, do errors discovered on a few cases from a small hospital (1 of 100, for example) have the same importance as a pattern of errors discovered in a large facility (20 of 1,200)? If the random sampling was stratified, the statisticians on the audit team should determine an overall estimated error rate by data item for the central registry as a whole.

Another question to be addressed is the distribution of the final report. Should it be published on the central registry's Web site? Who, besides the quality assurance team should see it? Certainly researchers using the data should be informed of any findings of concern, such as patterns of data errors or the accuracy of staging and treatment codes. Beyond internal dissemination of the report at the central registry, decisions should be made about submitting the report to the legislature or the central registry advisory board or others with vested interests in the quality of registry data.



Finally, the overall report needs an action plan. The action plan is based on interpretation of the analysis and the overall experiences of the auditors. For a reabstracting audit, did the final analysis indicate that only certain fields (or groups of fields like the staging section) were problem areas? If so, the action plan should describe training efforts, either for all data collectors in the state or for individual data collectors. Did the audit reveal that there were areas in the central registry's documentation that were not clear. If so, the action plan should describe the areas of the coding manual or the reportable list and reporting requirements that should be clarified. Did the audit indicate that the protocol was not clear enough to the participants? If so, revise the protocol now while it is still fresh in mind, so that the improvements won't be forgotten the next time a similar audit is done in the future.

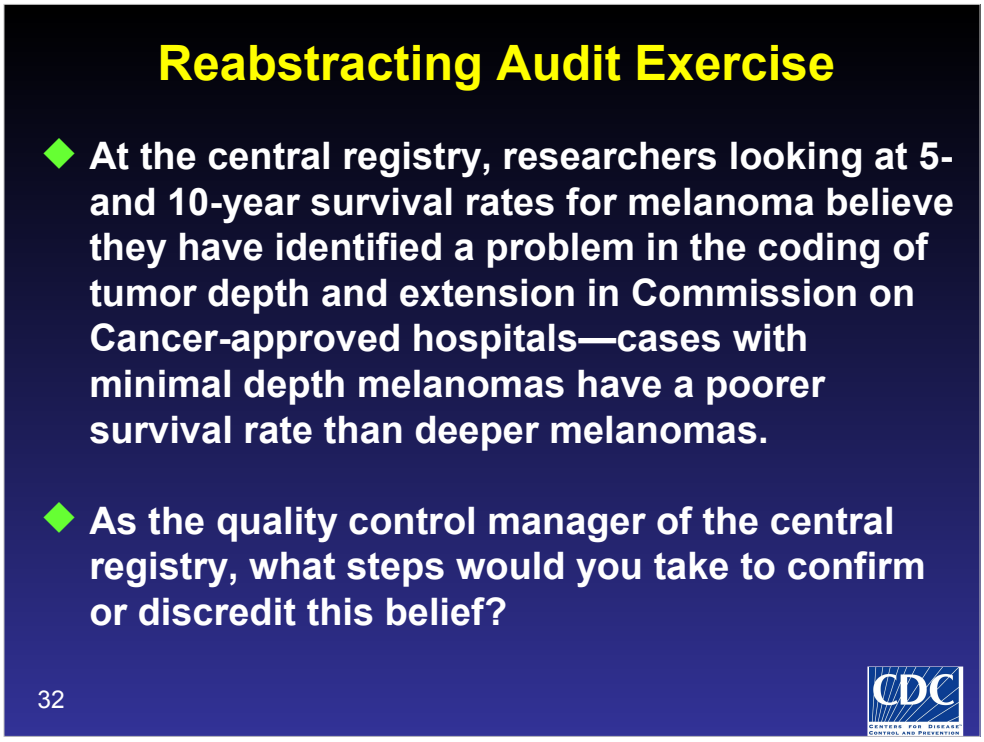
This reabstracting audit is complete. Now it's time to step back, take a deep breath, and begin the quality improvement cycle once again.



Before we finish, let's talk for a minute about recoding audits. Recoding audits are less expensive to conduct but cannot detect errors in abstracting. A recoding audit independently assigns codes on an abstract based on the supporting text documentation but without looking at the original codes. In essence, a recoding audit is an extension of the visual review process, but requires the auditor to code from the text without knowing what the original codes were. Like a reabstracting audit, any discrepancies must be reconciled with the original abstractor/coder. Then the results are tablulated and used as the basis for planning training. True errors remaining after reconciliation must be corrected in the database.

This type of study can be conducted at the central office, avoiding the need for travel, and is useful in training new coders. Concordance rates should be higher for recoding audits than for reabstracting audits because the source information is more limited. Recoding audits place emphasis on the quality of the text justification submitted with the abstract. If an abstractor does not do well on a recoding audit, more training is needed on how to write informative text, in addition to training on how to code medical information.

Recoding audits are less expensive to conduct because no travel is involved, although most other facets of a formal audit are still necessary, including development of audit software, writing a formal protocol, random sampling of cases, and completing follow-up documentation.



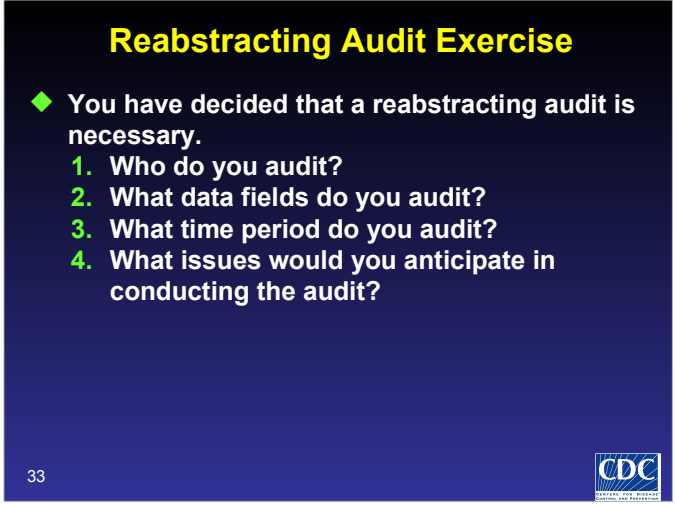
Let's try to focus what we've discussed on an issue raised by researchers using central registry data.

[read the slide to explain the scenario]

What steps would you take?

Possible answers (one or more):

1. Review a sample of cased abstracted prior to the implementation of Collaborative Staging in 2004, when all standards setters came to agreement on the definition of tumor size and depth of invasion.
2. Compare the distribution of tumor sizes and depth of invasion for COC-approved facilities and data collected by central registry abstractors.
3. Analyze cases by stage at diagnosis and survival. (If stage is increasing, survival should be decreasing.)
4. Ask the facilities to do a spot check of the data that was coded in tumor size prior to 2004.
5. Conduct a formal reabstracting audit to document data quality for the researcher. (This might be the final action because of the expense involved.)



Possible answers:

1A. Audit the COC-approved facilities. Prior to 2004, the *ROADS/FORDS* manual and SEER had different rules for coding depth of invasion in the tumor size field. *ROADS/FORDS* said to code tumor depth in millimeters; SEER said to code depth in hundredths of millimeters.

1B. Audit a random sample of melanoma cases for comparison from facilities unaffected by *ROADS/FORDS* rules, in other words, facilities not involved in COC approval.

2. Audit the tumor size field prior to 2004, or site-specific factor 1 in collaborative staging. Correlate tumor size with extension or TNM T category or the Extension field in EOD if available. In addition, treatment (surgery) and follow-up information would be useful.

3A. Audit cases diagnosed prior to 2004 when the rules became consistent. The research is basing the suspicion of inaccurate data on data used to calculate 5- and 10-year survivals. Although there may be problems accessing these older medical records, it would be important to document the tumor depth as reported in the medical record and compare that to how the tumor depth was coded on the abstract. It might be possible to look only at the case abstracts if the central registry required text documentation of codes prior to 2004, but this wouldn't be as accurate as comparing to the source documents.

3B. It might be possible to perform a very limited reabstracting audit using only pathology reports to determine depth of invasion, but the data could not be correlated with survival rates based on follow-up information from the medical record.

4A. Access to older records would probably be the biggest issue. Documents more than 5 years old might be on microfiche or microfilm, which is not pleasant to review for long periods of time.

4B. The older records may be less complete than what is required now. Summary stage is not conducive to detailed analysis of melanoma cases, and many central registries did not and do not require reporting of TNM data.

4C. The registrar at the audited facilities may have reported cases to the central registry using *ROADS/FORDS* rules, and the central registry quality control staff may not have caught the discrepancy at the time the case was added to the central registry database.

Resources

1. Dryden M and Brogan K. Quality Control. Chapter 20 in Menck H et al., *Central Cancer Registries: Design, Management and Use, second edition*. Kendall Hunt Publishing Co., 2007.
2. NPCR Educational Materials for Cancer Registrars
   - Volume 6: Audits: Casefinding and Reabstracting: Procedures for Central Registries
3. Unpublished materials provided by National Program of Cancer Registries

The findings and conclusions in this presentation are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

For information about CDC's
Cancer Prevention and Control Programs
and the
National Program of Cancer Registries

Please visit www.cdc.gov/cancer/npcr